

AUTOMATIC CONSTRUCTION OF MULTI-DOCUMENT MULTIMEDIA SUMMARIES

Mots clés :

- **Directeur de thèse** : bernard MERIALDO
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire de recherche d'EURECOM
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

With the increase of video quantity on the Internet, multimedia information processing has been a focused topic in the recent years. Among the techniques for multimedia information processing, video summarization has become an important tool. Some successful approaches have been proposed by the researchers in multimedia community. In this thesis, we propose our novel video summarization algorithm, Video-MMR (Video Maximal Marginal Relevance) based on visual information by mimicking MMR (Maximal Marginal Relevance) in text summarization. Video-MMR is a generic algorithm regardless of the video genre and suitable for summarizing both a single video and a set of videos. Besides Video-MMR as our basis, we also develop our approaches as following: (1) Since in a video, visual information is only one of several cues, more variants of Video-MMR using multimedia cues are proposed by exploiting multimedia information such as text or audio. We extend Video-MMR to AV-MMR (Audio Video Maximal Marginal Relevance), Balanced AV-MMR, OB-MMR (Optimized Balanced Audio Video Maximal Marginal Relevance) and TV-MMR (Text Video Maximal Marginal Relevance). These multimedia MMR algorithms are generic algorithms which outperform Video-MMR if we take into account the text and audio information in the video. (2) Exploiting more multimedia cues is only one way to improve Video-MMR, and visual information is always the most important among visual, acoustic and textual information. So in another way we overcome limits of Video-MMR and propose a refinement, Video-MMR2, by only exploiting visual information. (3) In addition to the summarization algorithms, we also optimize the presentation measure of video summaries, otherwise a good summary can be corrupted by a bad presentation. So we try optimizing a static summary containing keyframes and keywords by suggesting the number of frames and text grams, and dynamic summary composed of video segments by optimizing average duration of segments. (4) In the domain of video summarization, we need an evaluation measure for new approaches. Many current measures are based on human assessment, and the automatic evaluation method for video summaries is still an open problem. In this thesis we propose an approach, VERT (Video Evaluation by Relevant Threshold) mimicking the evaluation measures BLEU and ROUGE in text community to facilitate the automatic evaluation procedure with the help of only a few human assessments. We describe the details of all the approaches and present experimental results. Therefore, a framework on video summarization, including an algorithm of video summarization using visual cue, its variants exploiting more multimedia cues, an optimization measure of summary presentation, and a new evaluation method of video summaries, is proposed and allows us to manage and browse multiple videos more efficiently.